DOCUMENT RESUME

ED 316 892 CS 507 043

AUTHOR Dowling, Ralph E.
TITLE Reflective Judgment in Debate: Or, The End of "Critical Thinking" as the Goal of Educational Debate.
PUB DATE 1 Sep 89
NOTE 20p.; Paper presented at the Annual Meeting of the Western Speech Communication Association (Sacramento, CA, February 16-20, 1990).
PUB TYPE Viewpoints (120) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Cognitive Development; Critical Thinking; *Debate; Higher Education; *Models; Persuasive Discourse; Problem Solving; Research Methodology
IDENTIFIERS Competitive Argument; Debate Theory; Educational Issues; Piaget (Jean); *Reflective Judgment Model; Research Priorities; Watson Glaser Critical Thinking Appraisal

ABSTRACT

Recurring pressures for fiscal restraint threaten the existence of educational programs, such as competitive debate, which are not publicly perceived to produce worthwhile outcomes. Since debate is misunderstood and expensive, its advocates must be prepared to provide solid evidence of its benefits. Unfortunately, methodological weaknesses in debate research have prevented the accumulation of such evidence. The atheoretical nature of the critical thinking concept, particularly as measured in existing debate studies, exacerbates this problem. The critical thinking measure now used, the Watson-Glaser Critical Thinking Appraisal (WGCTA) offers a limited range of scores for assessing college students' critical thinking abilities, and the choice of behavior measured is not grounded in any particular theoretical formulation of human cognition. A promising new approach from the field of cognitive development--the reflective judgment model--provides an alternative that may remedy these deficiencies and secure a promising future for debate in higher education. The model has a clear foundation in cognitive developmental theory, philosophy, definitions, and theorization, and has been validated by a growing body of empirical data. It suggests that the skills it measures (which resemble those practiced in academic debate) are teachable. The model deals with problem-solving skills most useful to the real world and which develop in late adolescence and young adulthood--the age of interest to debate educators. (A 44-item bibliography is attached.) (SG)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED316892

REFLECTIVE JUDGMENT IN DEBATE:

OR, THE END OF "CRITICAL THINKING"

AS THE GOAL OF EDUCATIONAL DEBATE

Submitted as a Competitive Paper to:

The Western Forensic Association

for its Annual Meeting in Conjunction
With the Western Speech Communication Association

February 16-20, 1990

by

Ralph E. Dowling, Ph.D.
Department of Speech Communication
Ball State University
Muncie, IN 47306
(317) 285-1882
BITNET Address:
00REDOWLI@BSUVAX1

September 1, 1989

CS507043

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Ralph E. Dowling

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

ABSTRACT

Recurring pressures for fiscal restraint threaten the existence of educational programs not publicly perceived to produce worthwhile outcomes. Since debate is misunderstood and expensive, its advocates must be prepared to provide solid evidence of its benefits. Unfortunately, methodological weaknesses in debate research have prevented accumulation of such evidence. The atheoretical nature of the critical thinking concept, particularly as measured in debate studies, exacerbates this problem. A promising new approach from the field of cognitive development--the reflective judgment model--provides an alternative that may remedy these deficiencies and secure a promising future for debate in higher education. This paper explores the weaknesses of existing debate research, the theoretical and operational inadequacies of the "critical thinking" approach, and the nature of the reflective judgment paradigm. It then suggests the most appropriate course for future research into debate and reflective judgment.

REFLECTIVE JUDGMENT IN DEBATE:
OR, THE END OF "CRITICAL THINKING" AS THE GOAL OF EDUCATIONAL DEBATE

In *Magister Ludi*, Herman Hesse writes of a fictional academic enclave, called Castalia, in which students are trained to play an educational game known as the Glass Bead Game. The game pits students against one another in exercises involving the gathering and artistic synthesis of ideas from different disciplines into elegant configurations. Champions of the game are heroes to some, though not a few academicians and members of the surrounding community find the game sterile and pointless. As Hesse wrote:

> Many in Castalia, and some in the rest of the country outside the Province, regarded this elite as the ultimate flower of Castalian tradition, the cream of an exclusive intellectual aristocracy, and a good many youths dreamed for years of some day belonging to it themselves. To others, however, this elect circle of candidates for the higher reaches in the hierarchy of the Glass Bead Game seemed odious and debased, a clique of haughty idlers, brilliant but spoiled geniuses who lacked all feeling for life and reality, an arrogant and fundamentally parasitic company of dandies and climbers who had made a silly game, a sterile self-indulgence of the mind, their vocation and the content of their life. (118-19)

I trust that I need not belabor the parallels between Hesse's game and contemporary academic debate. Since those who find debate sterile may hold positions allowing them to influence the financing of competitive debate, this perception is troubling.
Altering this perception is not easy.
Hesse described what could and could not be done to assure the Game's future.

> We cannot do it by compulsory means, say by making the Glass Bead Game an official subject in the lower schools, nor can we do it by invoking what our predecessors meant this Game to be. We can prove only that our Game and we ourselves are indispensable by keeping the Game ever at the center of our entire cultural life, by incorporating into it each new achievement, each new approach, and each new complex of problems from the scholarly disciplines. We must shape and cultivate our universality, our noble and perilous sport with the idea of unity, endowing it with such perennial freshness and loveliness, such persuasiveness and charm, that even the soberest researcher and most diligent specialist will ever and again feel its message, its temptation and allure. (214)

Debate educators often and quite naturally feel that debate's benefits require little proof. As Hesse's Game Masters saw it, "Every day we witness the phenomenon: young elite pupils who have signed up for their Game course without any special ardor . . . are suddenly seized by the spirit of the Game, by its intellectual potentialities, its

venerable traditions, its soul-stirring forces, and become our passionate adherents and partisans" (215). As McBath observed of debate educators, "some will assert that the benefits are so self-evident as to make justification superfluous. A few may even resist the call to professional introspection. Most educators, however, will be drawn to the challenge of taking stock of their profession" ("Toward" 5).

Forensics educators to date have too seldom seen the need to justify "the existence of our elite" of debaters. As Hesse's game officials did, they have come to feel that the elite of debaters "are more than a reservoir of talented and experienced players from which we fill our vacancies and draw our successors." Rather, the elite debaters are seen as "an end in themselves" (215-16).

The rest of the world, however, may view debate as a luxury and its participants as idlers. If debate is to continue receiving adequate funding, it must be justified to others in terms they will understand and accept. Educational administrators and policy makers are not unsusceptible to trends and buzz-words, and the current fads in education seem to be assessment and critical thinking. These, then, are the terms in which debate educators must demonstrate debate's value.

Undoubtedly, many deplore and/or fear the trend to require quantification of educational outcomes and believe critical thinking an impractical goal. But, as this is the perspective of significant others, forensics educators must recognize the desirability of justifying debate budgets in these terms, and proceed to do so.

Assessment in Education and Debate

"Educational assessment," "value-added testing," and "program evaluation" are buzz-words that nevertheless signify genuine trends in American education. Faced by increasing economic and technical competition from abroad and by a sobering sense of fiscal restraint in the electorate, American educators are making and answering calls for assessment of the contributions education makes to students' lives. Such calls for assessment frequently are accompanied by specific statements of areas in which improvement and assessment are sought. The current fad appears to be in the area of critical thinking. According to O'Keefe,

> Ernest Boyer gives our secondary schools a mixed report card. The top 10 to 15 percent of American students receive an outstanding education which includes learning to remember and respond as well as to think creatively and critically. Of the remaining students, those who get something out of high school (around 60 percent) receive little in the way of intellectual challenge. The serious problem teachers face is one of encouraging all students, not just the elite, to move beyond rote memorization and recall and into more analytical and probing thinking skills. (2)

One of the organizations calling for assessment and reform in teaching critical thinking is the National Assessment of Educational Progress (NAEP). As Lawrence reports, NAEP has "called for major changes in how and what American students are taught,

based on 20 years of evaluations that indicate a disturbing lack of high-level achievement across the board". Accordingly, "It is apparent that fundamental changes may be needed to help American schoolchildren develop both content knowledge and the ability to reason effectively about what they know--skills that are essential if they are to take an intelligent part in the worlds of life and work" (Lawrence).

I do not contend that "critical thinking" is clearly defined and understood. In fact, one purpose of this essay is to clarify a theoretical approach to "critical thinking" that avoids the current ambiguity in the use of the term, while still allowing debate educators to hitch a ride on the "critical thinking" bandwagon. Rohler notes of the term:

> Critical thinking is the latest educational shibboleth. When a phrase has achieved the status of a slogan that is endlessly repeated by people with widely varying orientations, . . . it lacks a precise meaning. Critical thinking has become a God phrase--a glittering generality whose very ambiguity allows it to be embraced by any educational theorist or reformer. (1)

Because of its high per capita cost, debate seems especially vulnerable to demands for accountability and threats of fiscal homicide. And, unfortunately, existing research in debate has not been of the high quality and quantity to make a convincing case that debate is worth the cost because it produces great improvements in debaters' thinking skills. One of the main purposes of this paper is to show how future research into the cognitive educational benefits of debate can be improved by the adoption of a new conceptual approach to "critical thinking." First, I will review the existing research.

The Inadequacy of Current Debate Research

The weaknesses of the research claiming cognitive benefits from debate experience are legion. Much of the evidence is anecdotal. And, much of it reports data gathered from self-selected survey respondents (those choosing to return surveys) from among the already self-selected population of former debate participants (those who chose to debate).

Other studies are weakened by reliance on testing self-selected participants, or self-selected matched groups of participants and non-participants. The difficulty of pre- and post-testing anything but self-selectors seems as inherent to debate research as to research into the health effects of smoking. Hence, some of the more sophisticated design features found in the latter--such as matching of control and experimental group members on all possible relevant variables--may be required to establish persuasively the educational value of debate.

Even without these weaknesses, debate research would be flawed. The small number of studies, small sample sizes, and ambivalent findings provide little confidence that any benefits of debate training have been demonstrated.

Finally, existing research is almost exclusively based upon the theoretically questionable "critical thinking" concept and its problematic operationalization via

the Watson-Glaser Critical Thinking Appraisal (WGCTA; see Watson and Glaser). I will now discuss these issues in some detail.

I will not critique at length the reports of purely anecdotal statements by prominent ex-debaters praising the importance of debate training to their success in life. Such statements are too easily labeled as biased and can be discounted easily because of small samples, lack of systematic data collection, error built into the operational measure, and self-selectivity of both the ex-debaters available for testimony and those willing to respond.

A close relation to purely anecdotal research is survey research on former debaters. Matlon and Keele surveyed former participants in the National Debate Tournament and found the vast majority had graduate degrees, the most common careers were in education and law, and about 10% were working in the legislative and executive branches of government. Respondents also reported having learned critical thinking through their debate experience. This study is vulnerable to the aforementioned indictment of self-selection, as well as to the legitimate question of whether delayed self-reports of non-experts are a valid and reliable measure of improvements in critical thinking and/or their relationship to debate experience. The same two indictments apply to Arnold's survey of attorneys.

Some more sophisticated studies have attempted to measure increases in critical thinking resulting from debate training and/or experience (Howell; Williams; Beckman; Jackson; Cross; Colbert). With one exception (Brembeck), these studies used self-selecting experimental groups rather than randomly assigning subjects to experimental (debate) and control (no debate) groups.

Some of the studies found significantly higher pre-test scores on critical thinking for debaters than controls (Howell; Williams; Cross; Colbert). This suggests the importance of controlling for self-selection. Other findings of significant personality and developmental differences between debaters and non-debaters further suggest reason for concern. Matching students on some other traits (e.g. I.Q., age, grade point average) cannot compensate for self-selection until we can confidently determine what all of the relevant variables are.

Unfortunately, unless we wish to measure the effects of only very limited debate training/experience, random assignment to experimental and control groups may never be possible. And, I would argue, studies of the effects of long-term exposure to debate are of greater interest to educators. Hence, matching of cohorts on relevant variables may be the only course of action for debate researchers to take in establishing the benefits of debate.

Follert and Colbert's meta-analysis of the studies by Brembeck, Howell, Jackson, Cross, and Williams offers a strong challenge to the notion that the link between debate experience and improved critical thinking has been established firmly. In a meta-analysis of these five studies, Follert and Colbert found there was an 88% chance that the improvements in critical thinking discovered in these studies could be accounted for by chance and concluded

that their meta-analysis casts "substantial doubt on the claimed relationship between debate training and critical thinking improvement" (10).

All of the empirical studies of debate and critical thinking, then, suffer from at least one serious methodological flaw. And, the combination of several studies has been shown to offer little evidence for debate's benefits. In addition to Follert and Colbert's critical meta-analysis, criticisms of this research have been offered by McGlone and Andersen.

As Andersen has observed, the critical thinking measure used in all of these studies is the Watson-Glaser Critical Thinking Appraisal (WGCTA). In the next section I will argue that the use of the WGCTA may fatally flaw these studies.

Weaknesses in the WGCTA

The WGCTA is flawed for several reasons. First, it offers a limited range of scores for assessing the critical thinking abilities of college students. Second, while it measures specific and important behaviors (Helmstadter 1215; Woehlke 684), the choice of behaviors measured is not grounded in any particular theoretical formulation of human cognition. And, finally, the WGCTA is not consistent with what I feel to be the most promising theoretical approach to cognitive development in young adults--King and Kitchener's reflective judgment paradigm (Kitchener and Kitchener; Kitchener, *Intellectual*; King).

The WGCTA's manual suggests its utility in measuring "gains in critical thinking abilities resulting from instructional programs in schools, colleges, and business and industrial settings" (Watson and Glaser 2). In reviewing the WGCTA, Crites has concluded that "there appears to be an insufficient range on the test, however, for college students, particularly those in their last year." According to Crites, this limited range "raise[s] a question about whether the Watson-Glaser is appropriate for use at the higher educational levels, as the Manual implies."

A greater range of scores may be difficult to arrange by modifying the WGCTA because the abilities it tests are so commonly found in older college students. This may be because most older students have achieved what Piaget called the "formal operational" stage of cognitive development and because achieving this stage is a sufficient or nearly sufficient condition for mastery of the five kinds of skills measured by the WGCTA (Inhelder and Piaget; Watson and Glaser 1-2). That is, the WGCTA may measure kinds of skills not related to training beyond basic literacy. If so, a modified WGCTA will not resolve this unless it contains items designed to measure reasoning skills different *in kind* from those currently measured. These new items would have to measure skills less commonly found among persons who have achieved the formal operational stage of development. Any research based on the WGCTA which involves college students may be hopelessly confounded unless the relationships between cognitive development, age, and critical thinking are first resolved through theoretical and empirical inquiry.

The WGCTA apparently does not claim to have a theoretical foundation in cognitive development or elsewhere. Hence, it is difficult for researchers using

the WGCTA to determine, much less control/account for, relevant variables other than the independent. Without a theoretical foundation to predict what other variables might be relevant, matching and stratification cannot be used to control them. Hence, WGCTA-based research cannot realistically hope to avoid likely confounding variables or eliminate likely competing hypotheses.

Crucially, critical thinking researchers lack a theoretical foundation for the claim that critical thinking--as measured by the WGCTA--is a teachable skill. The five skills reportedly measured could easily represent innate traits or the achievement of new cognitive structures rather than teachable skills. Eliminating the weaknesses in the WGCTA, then, would require *post hoc* theorizing to justify its inclusion of the five skills it claims to measure. And, for college subjects, it will require in-kind modifications of the test items.

I will argue, however, that it is preferable to abandon the morass created by the atheoretical WGCTA and to adopt an approach to "critical thinking" more consistent with current trends in the study of cognitive development. This "reflective judgment" approach has a theoretical base that allows for the control of relevant variables through matching or stratification. It also offers a useful range of scores for college-age subjects. And, perhaps most importantly, the reflective judgment approach may provide a theoretically sound measure of the kinds of skills actually enhanced by sustained debate training and experience.

Kitchener and King's reflective judgment concept stems from a critique of Piagetian developmental theory and is based on the philosophical system of Karl Popper. The general superiority of the paradigm is argued as stemming from its improvements on the work of Piaget, while its specific superiority for debate research is that it deals with and is measured by the kinds of intellectual problems found in debate.

Reflective Judgment, Piaget, and Critical Thinking

The most widely accepted theory of human cognitive development is that of Jean Piaget. Piaget posits four stages of cognitive development through which all humans pass between birth and age sixteen (Inhelder and Piaget). This theory holds that at each stage of development, humans develop new cognitive structures. These new structures cannot emerge unless those of the previous stage have emerged. And, these new structures are responsible for the new kinds of cognitive skills that characterize each stage.

The first stage--the *sensorimotor period*--lasts from birth to age two. In this period the neonate develops reflexes and random movements; moves on to develop the important abilities to accommodate and assimilate; and discovers cause and effect, the idea of permanence, goal setting, imitation of others, experimentation, memory, thought, problem solving, and a self-concept. From two to six years of age the child experiences the *preoperational period*. Here the child begins imaginative thinking and develops subjective logic. Vocabulary increases from two hundred to two thousand words, and the ability to interpret language less literally and in a more sophisticated fashion develops.

The *concrete operational period* takes the child from age six to age twelve. During this stage the child begins to understand conservation, reversibility, and sets. The child also becomes able to decenter when reasoning, replaces imagination with reliance on literal facts, desires simplicity and order, and is better at visual than verbal problems. From ages eleven to sixteen the adolescent goes through the fourth and final stage--*formal operations*. The formal operational thinker can use formal logic in the propositional hypothetico-deductive method familiar to science, learns to reason abstractly, and matures personally, socially, and physically (Owens, Blount and Moscow, 34-43).

The Piagetian system, then, ends with adolescence and fails to substantively acknowledge any adult development or differences between the ways adolescents and adults think. Because the Piagetian system sees development as complete during adolescence, it retards research into adult thought. Reflective judgment researchers believe this to be a major theoretical weakness that leaves educators unprepared for the challenges of educating adults and measuring the results of that education. Reflective judgment researchers try to remedy this through a program of theorization and research.

Research in the Piagetian tradition measures the developmental stage of subjects by observing performance on tasks appropriate to each stage. Kitchener argues that the nature of these tests further inhibits discovery of differences between adolescent and adult thinking. The traditional Piaget tasks are what are called "puzzles" or "well-structured problems." The distinguishing feature of a puzzle is that "all the elements necessary for a solution are knowable and known, and there is an effective procedure for solving it" ("Cognition" 224).

The importance of this use of puzzles is revealed by Kitchener's three level model of cognition ("Cognition" 223-25). The first level, *cognition*, refers to simple cognitive functions such as "computing, memorizing, reading, perceiving, acquiring language, etc." The second level, or *metacognition* includes "processes which are invoked to monitor cognitive processes when an individual is engaged in level 1 cognitive tasks." These metacognitions are a necessary part of solving puzzles and well-structured problems, and include our knowledge of problem-solving techniques and strategies, how to use them, and their success or failure. These metacognitive processes are accounted for in Piagetian development and are sufficient to solve the puzzles with which stages are measured.

But life too rarely presents us with such problems. "The problems most often encountered in the real world . . . are of the ill-structured variety." Ill-structured problems have no single unequivocal solution that can be reached simply by using the proper cognitive and metacognitive process. For such ill-structured problems (of which good debate propositions are prime examples), "evidence, expert opinion, reason, and argument can be brought to bear on the issues, but no effective procedure . . . can guarantee a correct or absolute solution" (Kitchener, "Cognition" 224-25).

Adults' ability to deal with such ill-structured problems requires meta-metacognitions, or *epistemic cognitions*.

This third level of the reflective judgment model includes cognitions used "to monitor the epistemic nature of problems and the truth value of alternative solutions." Epistemic cognition includes "the individual's knowledge about the limits of knowing, the certainty of knowing, and the criteria for knowing." Finally, epistemic cognition "includes the strategies used to identify and choose between the form of solution required for different problem types" (Kitchener, "Cognition" 225-26).

Kitchener asks researchers on adult reasoning to "recognize the tie between ill-structured problems and epistemic cognition" and she and her colleagues have begun a program of research to do so. Debate educators should be interested in this tie as well, because, as Kitchener notes, "issues of jurisprudence, public policy, . . . [and] philosophy . . . are all areas in which epistemic assumptions are critical because they are all concerned with ill-structured problems" ("Cognition" 230-31). These are the very issues with which debate normally deals.

The WGCTA--which contains only well-structured problems--can measure only cognitive and metacognitive functioning. Since debate provides extended and repeated practice in resolving ill-structured problems of the sort identified by Kitchener, debate might most logically be said to be developing the epistemic cognitive level of adult reasoning not measured by the WGCTA. Since an improved test of debate's contribution to cognitive functioning must measure the epistemic cognitive level, and the reflective judgment perspective promises to do just this, "critical thinking" and the WGCTA should be abandoned in favor of reflective judgment and its measure.

The reflective judgment paradigm suggests that persons' epistemic cognitive functioning may be at one of seven stages or levels (see Kitchener and King; Welfel and Davison 210-11). Stages one and two involve absolutist thinking in which legitimate authorities who are in possession of manifest truth are the only justification needed for beliefs. The realization that legitimate authorities sometimes disagree leads to stage three in which absolute knowledge comes to be seen as existing in particular fields. Some fields have certain truth, and in these areas legitimate authorities still possess truth and are the ultimate justification for beliefs. In other fields we must wait for truth to become known. In the interim, any opinion will do because no belief can be verified or disconfirmed. At this stage evidence is viewed quantitatively rather than qualitatively, rather like the novice debater's claim to have won an argument because "They have only one card on this and we have three!"

The dissonance caused in educational settings by the holding of unjustified beliefs motivates movement to stage four. In stage four, the perception that uncertainty is temporary is replaced by the skeptic's realization that uncertainty is inherent to knowledge. The adult comes to use self rather than authority as the measure of personal truths. Lacking methods and criteria for discovering truth, the stage four reflective thinker's truths are idiosyncratic and may be unreliable bases for action. Movement to stage five requires the individual to learn rules for evaluating arguments and evidence so that competing beliefs can be evaluated for relative strength. Stage

five thinkers, then, justify their beliefs by rules appropriate to the context of each belief, but lack universal means of integrating beliefs from different fields.

Within stage five adults compare various perspectives with their own experiences and other perspectives. This process leads to the emergence of stage six, in which adults learn to transcend individual frames of reference and to evaluate claims with the aid of principles of inquiry general enough to apply across the many frames of reference.

Stage six thinkers, then, see beliefs as justified and plausible within a context limited by person, case, time, place, etc. But, importantly, they do not see objective knowledge as a goal or standard for beliefs. Stage seven thinkers, though, recognize that some claims are more correct than others despite the inherent uncertainty of knowledge. These claims are more correct because they more closely resemble reality. Stage seven thinkers also show flexibility in evaluating beliefs from different domains and remain aware of the uncertainty of knowledge.

Readers conversant with Perry's theory of cognitive development or with Popper's philosophy should see their influence in the descriptions of these stages. All educators may notice that "the descriptions of college-educated persons in university mission statements closely parallel several important components in the higher stages of reflective judgment" (Welfel and Davison 210). And, debate educators should notice the affinity between the goals of debate training and the higher levels of reflective judgment. These higher levels "are characterized by a growing sophistication in the capacity to interpret evidence, in objectivity in viewpoint, and in a conscious understanding of the process of problem solving" (Welfel and Davison 210).

If reflective judgment so closely parallels the goals of debate training, then reflective judgment would make a superior dependent variable in studies assessing the cognitive outcomes of debate training. Hence, the reflective judgment paradigm may be debate's best hope of proving it produces desirable cognitive outcomes. The WGCTA makes no claim to measure the higher levels represented by epistemic cognitions, and if these are what debate really teaches, we should not be surprised at the mixed results of studies using the WGCTA.

The weaknesses of the WGCTA and of the critical thinking approach are resolved by the reflective judgment paradigm. First, the paradigm has a clear foundation in cognitive developmental theory, philosophy, definitions, and theorization; and has been validated by a growing body of empirical data. Second, it suggests that the skills it measures are teachable. Third, these skills certainly resemble those practiced in academic debate. And, fourth, the paradigm deals with problem-solving skills most useful to the real world and which develop in late adolescence and young adulthood--the ages of interest to debate educators.

Reflective Judgment Research

Relationship With Other Concepts

An important part of the reflective judgment research program has been to establish the relationship between

reflective judgment and related developmental/ educational outcomes and variables. Reflective judgment levels are measured by coding the transcripts of semi-structured interviews (Kitchener and King). These interviews elicit respondents' epistemic cognitions through a discussion of four ill-structured problems known as dilemmas. These dilemmas involve controversies over theories of the construction of the pyramids, objectivity in journalism, creation/evolution, and the safety of chemical food additives (Kitchener, *Intellectual*; King).

Scores on the Reflective Judgment Interview (RJI) have been compared to a number of other measures. For example, Welfel and Welfel and Davison found that scholastic ability (as measured by the Preliminary Scholastic Aptitude Test) did not account for differences in subjects' changes in RJI scores during their college years. The same scholars measured verbal ability (using Terman's Concept Mastery Test, or CMT) and found that verbal ability also could not account for changes in RJI scores during the college years. Kitchener found an overall correlation of .79 between RJI and CMT scores, but the correlation differed for different subgroups of her subjects, and was low and non-significant for some of her cohorts (*Intellectual*).

Brabeck matched subjects within one point on WGCTA scores and found that RJI scores increase with increased higher education even when WGCTA scores are constant. Further, she found that while high RJI subjects were uniformly high on WGCTA, those who scored high on WGCTA had highly variable RJI scores. This strongly suggests that reflective judgment and critical thinking are different concepts. Development of critical thinking skills, then, appears to be a necessary but not a sufficient condition for achievement of higher levels of reflective judgment.

Brabeck's findings are consistent with the theoretical descriptions of the levels of reflective judgment, because the model suggests the emergence of critical thinking skills at level five. The findings and the model also are consistent with Mines's finding that some of the critical thinking skills measured by the WGCTA and the Cornell Critical Thinking Test are reliable predictors of whether subjects have achieved a reflective judgment level above three.

Other research involving RJI supports reflective judgment scholars' contention that reflective judgment involves something above and beyond Piaget's fourth and final stage of cognitive development--the formal operational. King found highly significant differences ($p<.001$) in RJI scores for high school, undergraduate, and graduate students despite no significant differences in scores on Piagetian measures of formal operations. 92% of her subjects measured at the formal operational level.

Reflective judgment, then, appears to be a different concept than scholastic aptitude, verbal ability, critical thinking, and formal operations. And, it appears to be different from these concepts in ways consistent with the reflective judgment model of epistemic cognition. These issues are not resolved conclusively, and research in this area continues.

Validity/Reliability

Operationalizing variables raises issues of reliability and validity. The previously discussed findings of the relationship between RJI scores and other measures would appear to offer some evidence of construct validity, as would the finding that RJI scores increase with number of years in college. More research of the latter variety will be reported in the next section of this paper.

Welfel and Davison reported that 92% of subjects' RJI scores increased after three years of college. King et al. reported that 90% of subjects' RJI scores improved after two years of higher education. King reported an internal consistency reliability coefficient of .96 for the RJI. Welfel and Davison reported an overall reliability coefficient of .89 for the RJI.

Relationships With Independent Variables

A primary goal of reflective judgment researchers has been to determine the effects of higher education on reflective judgment levels. The potential significance of such research is inestimable. If reflective judgment can first be proven to be enhanced by higher education, follow-up studies on the relative effectiveness of different curricula, formats, teacher styles, admissions standards, etc. could prove invaluable in higher education's ability to meet its stated goals, silence its critics, and rally its supporters. The same can be said of the potential benefits of this line of inquiry for debate educators.

Many studies have indeed found that as the number of years of higher education increases, so do the reflective judgment levels of students (King; Kitchener, *Intellectual*; Strange; Welfel; Brabeck; Schmidt; Welfel and Davison). Further, Lawson found that graduate students had higher RJI scores than non-students matched on both scholastic aptitude and age.

Stemming as it does from a Piagetian foundation, the reflective judgment model naturally raises the question of whether reflective judgment levels represent teachable skills or structures obtainable only through maturation. Hence, many studies have looked at both age and education along with RJI scores. Strange and Shoff both found that age alone did not affect RJI scores when the education of subjects was held constant. These findings are not consistent with Lawson's discovery that older subjects had higher RJI scores, and not completely consistent with Schmidt's finding that for women, RJI scores increased with age. Schmidt also found, for her sample as a whole, that the combination of age and education had more impact on RJI scores than either age or education alone.

These ambivalent findings suggest that age and education both contribute to reflective judgment even within the narrow age and education ranges of the high school to young adult populations sampled in these studies. Despite the presence of numerous studies showing a positive correlation between education and RJI scores, this is an area in which further research clearly is needed. Readers should note, however, that only some of the studies showing the positive correlation have been cited here.

Many mental measurements show sex or gender differences. The RJI is no exception, with many studies finding males scoring higher (Kitchener; Strange; Shoff; Lawson; Mines). Not all studies, however, have found this difference. The author is unaware of any analyses of psychological gender's relationship to RJI scores.

Summary

Significantly, then, reflective judgment appears to consist of skills teachable through higher education. Debate educators might do well to test whether their specific contribution to reflective judgment can be measured. Such specific studies of effects on RJI scores have been few and not promising. Sakalys found, not surprisingly, that an undergraduate research course was insufficient to increase the RJI scores of nursing students. Welfel and Welfel and Davison found that academic major appeared to make no difference in RJI score improvement over three years.

Proposed Research in Debate

A particularly promising niche remains open for debate to prove its unique worth in a program of higher education. Many reflective judgment studies have found that college seniors and graduate students--despite improvement in reflective judgment ability during their education--rarely achieve the higher levels of reflective judgment. For example, Welfel and Davison measured entering freshmen and remeasured four years later and found no student above level five of reflective judgment, with the majority being at level four. Shoff reported similarly low scores among freshmen. If debate can help students achieve these higher levels of reflective judgment--and prove that it can do so--its future role in higher education would be secure. Even if debate only enhances the reflective judgment of an elite that begins with above average reflective judgment ability, it will have a proven educational benefit justifying its existence.

Even if the RJI rather than the WGCTA is used by debate researchers, however, the other methodological weaknesses discussed earlier in this paper also must be remedied. That is, randomly selected control groups, or sophisticated matching of debaters with non-debate cohorts will have to be used. Appropriate pre- and post-tests on RJI, and both longitudinal and cross-sectional analyses must be used.

Once the proper instrument and designs have been chosen, a number of interesting research questions suggest themselves. The most apparent and important is, "What effect does debate training have on RJI scores?" This question leads to others regarding the effects of different amounts and types of debate training. Not to slight our individual events counterparts, this also suggests that, if all forensics is to have an argumentative perspective, we also should measure the effects of individual events training on RJI scores (McBath, *Forensics as Communication*).

We might also be interested in the RJI scores of students attracted to competitive debate in comparison to those of students not attracted to debate. And then there is the question of the comparative RJI scores (or changes in same) of those who continue to debate

and those who choose to drop out of debate.

The relationship between RJI scores and competitive success in debate is a fascinating prospect for research, as is any interaction effects among RJI scores, competitive success, and drop-out rate. We might all benefit from studying the relationship between RJI levels and success in debate coaching or judging.

From a more general communication perspective, we might also be interested in studies of the relationship between reflective judgment and such variables as argumentativeness (Infante and Rancer), communication apprehension (McCroskey), cognitive complexity (Delia), etc. If such studies included both debaters and nondebaters, they could also provide data of particular interest to debate educators.

Such a program of research will be costly and time-consuming. The RJI is copyrighted and may be administered only by those certified to do so after paying for and receiving training. The interviews must be tape-recorded, transcribed, and scored by scorers certified after paying for and receiving training. The author has even had difficulty obtaining funding for even a pilot study using the RJI.

But, in terms familiar to my readers, the advantages of pursuing this line of research may well be worth the costs. In a world of fiscal restraint, educational accountability, and widespread ignorance of debate's contributions to participants, only convincing proof of debate's value can assure it of any future. And, the reflective judgment perspective may well provide this convincing proof--something that critical thinking studies have not provided and may never provide. The stakes are high and the challenge is before us.

Works Cited

Andersen, Kenneth E. "A Critical Review of the Behavioral Research in Argumentation and Forensics." Journal of the American Forensic Association 10.3 (1974): 147-55.

Arnold, William E. "Debate and the Lawyer." Journal of the American Forensic Association 3.1 (1966): 26-28.

Beckman, Vernon E. An Investigation of the Contributions to Critical Thinking Made by Courses in Argumentation and Discussion in Selected Colleges. Diss. U of Minnesota, 1957.

Brabeck, Mary Margaret Knutson. The Relationship Between Critical Thinking Skills and the Development of Reflective Judgment. Diss. U of Minnesota, 1980.

Brembeck, Winston L. "The Effects of a Course in Argumentation on Critical Thinking Ability." Speech Monographs 16.2 (1949): 177-89.

Colbert, Kent R. "The Effects of CEDA and NDT Debate Training on Critical Thinking Ability." Journal of the American Forensic Association 23.4 (1987): 194-201.

Crites, John O. "Review of the Watson-Glaser Critical Thinking Appraisal." The Seventh Mental Measurements Yearbook. Ed. Oscar K. Buros. Highland Park, NJ: Gryphon Press, 1972. 1214.

Cross, Gary P. The Effects of Belief Systems and the Amount of Debate Experience on the Acquisition of Critical Thinking. Diss. U of Utah, 1971.

Delia, Jesse G. "Constructivism and the Study of Human Communication." Quarterly Journal of Speech 63.1 (February 1977): 66-83.

Follert, Vincent F. and Kent R. Colbert. "An Analysis of the Research Concerning Debate Training and Critical Thinking Improvements." Paper presented to the Annual Meeting of the Speech Communication Association, Washington, D.C.,, November 10-13, 1983. (ERIC ED 238 058)

Helmstader, G. C. Journal of Educational Measurement 2 (Dec. 1965): 254-56.

Hesse, Herman. Magister Ludi. Trans. Richard Winston and Clara Winston. New York: Bantam, 1970.

Howell, William S. "The Effects of High School Debating on Critical Thinking." Speech Monographs 10 (1943): 96-103.

Infante, Dominic A. and Andrew S. Rancer. "A Conceptualization and Measure of Argumentativeness." Journal of Personality Assessment 46.1 (1982): 72-80.

Inhelder, Barbel and Jean Piaget. The Growth of Logical Thinking from Childhood to Adolescence. Trans. Anne Parsons and Stanley Seagrin. New York: Basic Books, 1958.

Jackson, Teddy Randolph. The Effects of Intercollegiate Debating on Critical Thinking Ability. Diss. U of Wisconsin, 1961.

King, Patricia Margaret Brown. The Development of Reflective Judgment and Formal Operational Thinking in Adolescents and Young Adults. Diss. U of Minnesota, 1977.

King, P. M., K. S. Kitchener, M. L. Davison, C. A Parker, and P. K. Wood. "The Justification of Beliefs in Young Adults: A Longitudinal Study." Human Development 26 (1983): 106-16.

Kitchener, Karen Strom. Intellectual Development in Late Adolescents and Young Adults: Reflective Judgment and Verbal Reasoning. Diss. U of Minnesota, 1977.

---. "Cognition, Metacognition and Epistemic Cognition: A Three-Level Model of Cognitive Processing." Human Development 26 (1983): 222-32.

Karen Strohm Kitchener and Richard F. Kitchener. "The Development of Natural Rationality: Can Formal Operations Account for It?" Contributions to Human Development 5 (1981): 160-81.

Karen Strohm Kitchener and Patricia M. King. "Reflective Judgment Scoring Rules." Unpublished manuscript. Revised May 1985.

Lawrence, Jill. "School Report Indicates Basics Better." Muncie Star 15 Feb. 1989: A13.

Lawson, Jane Marie. The Relationship Between Graduate Education and the Development of Reflective Judgment: A Function of Age or Educational Experience. Diss. U of Minnesota, 1980.

Matlon, Ronald J. and Lucy M. Keele. "A Survey of Participants in the National Debate Tournament, 1947- 1980." Journal of the American Forensic Association 20 (Spring 1984): 194-205.

McBath, James H. Ed. Forensics as Communication: The Argumentative Perspective. Skokie: National Textbook Company, 1975.

McBath, James H. "Toward a Rationale for Forensics." American Forensics in Perspective. Ed. Donn Parson. Annandale, VA: Speech Communication Association, 1984.

McCroskey, James C. "Oral Communication Apprehension: A Summary of Recent Theory and Research." Human Communication Research 4.1 (1977): 78-96.

McGlone, Edward L. "The Behavioral Effects of Forensics Participation." Journal of the American Forensic Association 10.3 (Winter 1974): 140-46.

Mines, Robert Alan. An Investigation of the Developmental Levels of Reflective Judgment and Associated Critical Thinking Skills. Diss. U of Iowa, 1980.

O'Keefe, Virginia P. Affecting Critical Thinking through Speech. Urbana, Illinois: ERIC Clearinghouse, 1986.

Owen, Steven, Parker Blount and Henry Moscow. Educational Psychology: An Introduction. Boston: Little, Brown, 1978.

Perry, William G. Jr. Forms of Intellectual and Ethical Development in the College Years. (1968; Cambridge: Harvard UP; New York: Holt, Rinehart and Winston, 1970.

Popper, Karl R. The Logic of Scientific Discovery. New York: Harper, 1968.

Rohler, Lloyd. "Developing Critical Thinking Skills in Upper Level Speech Communication Courses." Unpublished paper presented to the Speech Communication Association Annual Convention. Boston. 1987.

Sakalys, J. A. Effects of Research Methods Course on Nursing Students' Research Attitudes and Cognitive Development. Diss. U of Denver, 1984.

Schmidt, Janet Adele. The Intellectual Development of Traditionally and Nontraditionally Aged College Students: A Cross Sectional Study With Longitudinal Follow-Up. Diss. U of Minnesota, 1983.

Shoff, Susan Patigalia. The Significance of Age, Sex and Type of Education on the Development of Reasoning in Adults. Diss. U of Utah, 1979.

Strange, Charles Carney. Intellectual Development, Motive for Education and Learning Styles During the College Years: A Comparison of Adult and Traditional-Age College Students. Diss. U of Iowa, 1978.

Watson, Goodwin and Edward M. Glaser. Manual of Directions for Critical Thinking Appraisal Forms A and B. San Antonio: The Psychological Corporation, 1980

Welfel, Elizabeth Reynolds. The Development of Reflective Judgment: Its Relationship to Year in College, Academic Major, and Satisfaction With Major Among College Students. Diss. U of Minnesota, 1979.

Welfel, Elizabeth Reynolds and Mark L. Davison. "The Development of Reflective Judgment During the College Years: A 4-Year Longitudinal Study." Journal of College Student Personnel 27.3 (May 1986): 209-16.

Williams, Donald E. The Effects of Training in College Debating on Critical Thinking Ability. Unpublished master's thesis, Purdue University, 1951.

Woehlke, Paula Lee. "Watson-Glaser Critical Thinking Appraisal." Test Critiques. Ed. Daniel J. Keyser and Richard C. Sweetland. Kansas City: Test Corporation of America, 1985.